

Crystal structure of a double-stranded DNA containing a cisplatin interstrand cross-link at 1.63 Å resolution: hydration at the platinated site

Franck Coste, Jean-Marc Malinge, Laurence Serre, William Shepard¹, Michel Roth², Marc Leng and Charles Zelwer*

Centre de Biophysique Moléculaire, Centre de National de la Recherche Scientifique, affiliated to the Université d'Orléans, rue Charles Sadron, 45071 Orleans Cedex, France, ¹Laboratoire pour l'Utilisation du Rayonnement Electromagnetique, Université de Paris 11, 91405 Orsay Cedex, France and ²Institut de Biologie Structurale, 41 rue des Martyrs, 38027 Grenoble Cedex, France

Received January 15, 1999; Revised and Accepted February 18, 1999

NDB/PDB accession no. DDJ075

ABSTRACT

cis-diamminedichloroplatinum (II) (cisplatin) is a powerful anti-tumor drug whose target is cellular DNA. In the reaction between DNA and cisplatin, covalent intrastrand and interstrand cross-links (ICL) are formed. Two solution structures of the ICL have been published recently. In both models the double-helix is bent and unwound but with significantly different angle values. We solved the crystal structure at 100K of a double-stranded DNA decamer containing a single cisplatin ICL, using the anomalous scattering (MAD) of platinum as a unique source of phase information. We found 47° for double-helix bending and 70° for unwinding in agreement with previous electrophoretic assays. The crystals are stabilized by intermolecular contacts involving two cytosines extruded from the double-helix, one of which makes a triplet with a terminal G-C pair. The platinum coordination is nearly square and the platinum residue is embedded into a cage of nine water molecules linked to the cross-linked guanines, to the two ammine groups, and to the phosphodiester backbone through other water molecules. This water molecule organization is discussed in relation with the chemical stability of the ICL.

INTRODUCTION

cis-diamminedichloroplatinum (II) (cisplatin) is a widely used anti-tumor drug in cancer chemotherapy (1). Numerous studies suggest that the therapeutic effect of cisplatin is related to its ability to cross-link cellular DNA. Cisplatin reacts with the N7 position of guanine residues to form monofunctional adducts which close further into bifunctional intrastrand and interstrand cross-links (ICL). The major intrastrand cross-links are formed at d(GpG) and d(ApG) sites and represent ~90% of total platinum adducts (2–5). ICL are formed between two guanine residues on

opposite strands at d(GpC)-d(GpC) sites (6,7) and represent 5–10% of total platinum adducts. The relative contribution of the different types of DNA lesions to the pharmacological properties of cisplatin is still unknown.

The relevance of the study of cisplatin adducts-induced DNA distortions has been recently emphasized by the discovery of several proteins which specifically recognize the major intrastrand cross-links (8,9 and references herein). The distortions induced by the major intrastrand adducts have been characterized by combination of different technical approaches (10). The three-dimensional (3D) structure of short double-stranded DNA duplexes site-specifically modified by a cisplatin 1,2-d(GpG)-intrastrand cross-link has been studied by ¹H NMR (11–13) and by X-ray crystallography at 2.6 Å resolution (14). The main distortions induced by this adduct are a bend and unwinding of the double helix and an alteration of the size of the grooves and of the puckering of the sugars.

Although cisplatin intrastrand cross-links are generally considered as playing a major role in the biological properties of the drug, the contribution of ICL to the mechanism of action of cisplatin should be taken into consideration. The number of ICL induced by cisplatin has been correlated with the cytotoxicity of the drug (15). Resistance of cells to cisplatin has been associated with increased gene-specific DNA repair efficiency of ICL (16–18). ICL strongly inhibit DNA transcription elongation by RNA polymerase (19). The distortions induced by cisplatin ICL have been first characterized by means of different techniques such as chemical probes, footprinting and electrophoresis experiments (20–23). Recently, ICL structure has been studied by ¹H NMR with two double-stranded DNA decamers of different sequences referenced in the present report as NMR1 (23) and NMR2 (24). The two proposed structures for ICL revealed very unusual distortions with respect to B-DNA. The double helix is largely unwound, bent towards the minor groove and the C residues complementary to the cross-linked G residues are unpaired. However, the DNA bending angles and directions in the two NMR structures are quite different, as well as the helix unwinding

*To whom correspondence should be addressed. Tel: +33 2 38 25 78 03; Fax: +33 2 38 63 15 17; Email: zelwer@cns-orleans.fr

values. Besides, neither experimental evidence about the structure of the platinum residue in ICL has been so far obtained nor information about the water molecules organization. Numerous data underline the importance of well-ordered solvent molecules for the conformational equilibrium of DNA (25,26), for the interaction between drugs and DNA (27,28) and for the recognition of DNA by proteins (29,30).

In this work, we solved the first crystal structure of a double-stranded DNA decamer containing an ICL and of sequence: d(C₁C₂T₃C₄G₅*C₆T₇C₈T₉C₁₀)·d(G₁₁A₁₂G₁₃A₁₄G₁₅*C₁₆G₁₇-A₁₈G₁₉G₂₀), where G₅* and G₁₅* are guanine residues on opposite strands of DNA cross-linked by cisplatin at the N7 position. In addition to the determination of the detailed structural features induced in DNA by cisplatin ICL, the high resolution structure reveals the existence of an unexpected organization of ordered water molecules around the platinum residue. These results are discussed in the context of the particular chemical reactivity of the bonds within ICL.

MATERIALS AND METHODS

Platinated oligonucleotide preparation and crystallization

The unplatinated oligonucleotides were synthesized and purified by Eurogentec. Cisplatin was from Johnson Matthey. The decamer site-specifically modified by a unique cisplatin ICL was prepared in two steps and purified as previously described in detail (24).

Crystals were obtained by vapor diffusion at 4°C using the hanging drop method. The initial conditions in the drop were 82.5 μM DNA, 20 mM sodium cacodylate, 5 mM NaCl, 30 mM KCl, 2.5 mM spermine, 2% (v/v) 2-methyl-2,4-pentandiol (MPD), pH 6.0. The drops were equilibrated against a solution of 2.5% MPD. The first crystals appeared after 1 week and reached their final sizes (200 × 200 × 100 μm) during the next 2 or 3 weeks. For data collection, crystals were transferred into a buffer containing 40% (v/v) MPD (other concentrations unchanged) during 10 min and then flash-frozen and stored in liquid nitrogen.

MAD data collection, phase determination and model building

The platinated double-stranded decamer crystallized in space group C2 with one molecule per asymmetric unit. Attempts to solve the structure by molecular replacement with the NMR2 atomic model were unsuccessful. In a preliminary MAD experiment (data not shown), several flash-frozen crystals were transported and tested at the D2AM beam-line at the European Synchrotron Radiation Facility (ESRF; Grenoble, France). The best one showed diffraction data until 2.4 Å resolution. Diffraction images were collected from this crystal at the L_{III} absorption edge of Pt, at four wavelengths selected with the help of a fluorescence scan made on the crystal. One wavelength corresponded nearly to the minimum of f' , another to the maximum of f'' and the two others were taken at ~15 eV on each side of the absorption edge. The detector was the X-ray image intensifier CCD camera developed at ESRF (31). Because of lack of beam time, no special care was taken for orienting the unique axis b of the crystal unit cell parallel to the spindle axis, and as a consequence the data sets contained only very few Bijvoet pairs. After correction of the images for distortion (program *imac*; M.Roth, unpublished) and their scaling (program *Hisp*; M.Roth), the reflections were

integrated using *XDS* (32). The scaling of the intensities was refined and the partial structure factors of the anomalous scatterers calculated with the program *nYn* (M.Roth) according to the equations of Karle and Hendrickson (33). With these partial structure factors, we were able to find the platinum atom position from a Patterson map and obtain an electron density map interpretable for model building. The initial choice of the enantiomorph was correct according to the double helix geometry.

After optimization of the crystallization parameters, new MAD data were collected at four wavelengths (λ_1 = low energy remote, λ_2 = inflection point, λ_3 = peak, λ_4 = high energy remote) at a maximum resolution of 1.6 Å from a single frozen crystal at the DW21b beam-line from LURE (Orsay, France), using a MAR300 imaging plate detector. The wavelengths, near the L_{III}-shell edge of platinum, were chosen from an X-ray fluorescence spectrum of the frozen crystal. The four data sets were recorded one at a time. Because of its natural shape, the crystal oriented itself spontaneously in the drop before flash-freezing with its dyad axis close to the rotation axis, enabling us to record Bijvoet mates on nearby frames. The data from the different images and for each wavelength were processed using *MOSFLM* (34) for integration and scaled using *SCALA* (32). A summary of the data collection statistics is given in Table 1. For the λ_3 data (peak) and λ_4 (high energy remote), the R_{ano} is double the R_{sym} .

The position of the platinum peak was confirmed by the anomalous Patterson map computed with the λ_3 amplitudes and the dispersive Patterson map at 1.7 Å resolution. Determination and refinement of the structure factor phases was done with the program *SHARP* (35) based on maximum likelihood ranking, taking λ_1 amplitudes as native. High figure of merit (overall: 0.84, for the 1.74–1.63 Å resolution shell: 0.77) and phasing power values (between 5.0 and 8.6) were obtained after convergence. The map computed with the MAD phases and after solvent flattening (*SOLOMON*; 36) at 1.7 Å resolution is of remarkable quality (Fig. 1). The oxygen atoms from the phosphate groups have well defined bulges with distinct density maxima. The bond

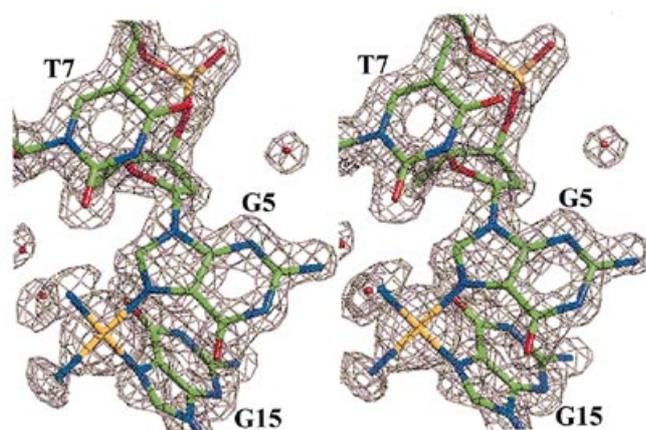


Figure 1. Initial weighted F_0 map computed with MAD phases and after solvent flattening, contoured at 1.5σ in the area of the platinum residue. The atomic model corresponds to the refined coordinates SHELL. Stick representation and map are from *TURBO-FRODO*. The atom color code is C, green; N, blue; O, red; P, orange; Pt, yellow.

Table 1. Statistics of data collection, phasing and structure refinements

Crystallographic data							
Space-group	C 2						
Cell dimensions	a = 42.89Å, b = 29.98Å, c = 46.57Å, β = 95.98°						
1 molecule / asymmetric unit							
MAD data							
Wavelength	Resolution limit(Å)	Number of measurements	Unique reflections	Completeness (%)	<I/σ(I)> (%)	R _{int} (%)	R _{meas} (%)
1.0906 (λ ₁) (1.82-1.72 Å)	1.72	51,152	6,203	94	13.5 (12.6)	3.5 (5.4)	4.1 (5.9)
1.0718 (λ ₂) (1.72-1.69 Å)	1.69	54,014	6,555	95	12.7 (8.5)	4.0 (8.3)	6.9 (12.3)
1.0711 (λ ₃) (1.72-1.69 Å)	1.69	77,145	6,601	97	12.9 (12.2)	4.1 (5.8)	8.9 (12.7)
1.0301 (λ ₄) (1.72-1.63 Å)	1.63	60,517	7,333	97	14.04 (11.1)	3.7 (6.3)	7.6 (10.2)
Phasing statistics							
		λ ₁	λ ₂	λ ₃	λ ₄		
R _{critis} acentrics anomalous		0.33	0.35	0.28	0.57		
R _{krout} acentrics anomalous		0.017	0.034	0.035	0.030		
Phasing power acentric anomalous		7.27	7.30	8.65	5.07		
Figure of merit		acentrics		centrics			
overall		0.84		0.77			
(1.74-1.63 Å) shell		0.77		0.68			
Refinement data							
Resolution (Å)	20-1.63						
No. reflections	7,333						
R-factor (%) (all data)	16.92						
Free R-factor (%)	20.32						
No. non-hydrogen atoms	499						
No. water molecules	92						
Average temperature factor (Å ²)							
nucleic acid atoms	18.07						
water molecules	30.25						
R.m.s. deviation from ideality							
bonds (Å)	0.007						
angles (Å)	0.016						
$R_{\text{sym}} = \frac{\sum_h \langle I_1(h) - \langle I_i(h) \rangle \rangle}{\sum_h \langle I_i(h) \rangle}$ where I _i (h) is the i th measurement of a group of equivalent reflections and <...> is the mean value of the quantity in brackets.							
R_{ano} is the measure of the mean relative anomalous difference between the Bijvoet pairs. $R_{\text{critis}} \text{ acentrics} = \text{r.m.s. lack of closure/r.m.s. anomalous difference.}$ $R_{\text{krout}} \text{ acentrics} = \frac{\sum F_{\text{PH(obs)}} - F_{\text{PH(calc)}} }{\sum F_{\text{PH(obs)}}}$							
Phasing power = $\langle F_H \rangle$ / r.m.s. lack of closure.							
$R \text{-factor} = \frac{\sum (F_{\text{obs}} - k F_{\text{calc}})}{\sum F_{\text{obs}} }$ Free R-factor is the R-factor obtained for a test set of reflections (5% of the diffraction data).							

angles are clearly defined for all non-hydrogen atoms of the molecule, and many well-ordered water molecules are identifiable from the solvent flattened map. However these water molecules were not included in the model at the early stages of the atomic model refinement. The atomic model was built with *TURBO-FRODO* (v.5.5) (37).

Model refinement

Model refinement was done against the structure factor magnitudes from the λ₄ data (1.63 Å resolution). The refinement program was *SHELX* (38) with stereochemical restraints. Bond distances and angles theoretical values were used as observations with weights equal to the observed structure factor magnitudes. Periodically, the model was visually inspected to track the process of the refinement and to manually adjust parts of the model. After convergence was obtained, water molecules were identified automatically on (2F_o-F_c) and (F_o-F_c) maps using *ARP* (39) and

then a new set of refinement cycles with *SHELX* was completed. This process was repeated several times. Water molecules detected by *ARP* but which did not give densities above 1σ on a (2F_o-F_c) map were removed from the model. These water molecules were checked for hydrogen bond linkage consistency (one to four neighboring hydrogen bond donors or acceptors at 2.4–3.2 Å distance). Individual isotropic B-factors were used throughout the initial refinement cycles. For the last cycles of refinement, anisotropic temperature factors were refined for atoms of the platinum residue and for the phosphorus atoms from the backbone. Two sets of refinement cycles have been done in parallel, one with geometrical restraints applied to the platinum residue distances and angles (*SHEL1*) and the other without these restraints (*SHEL2*). The differences between the two resulting conformations of the platinum residue are small but probably relevant. Indeed, the Pt-N bonds distances in the platinum residue being ~2 Å, the corresponding atoms are resolved at 1.63 Å, the

resolution of the data (Fig. 1). The *R*-factor of 32% for the initial model dropped to 17% after refinement by *SHELX*, (*R*-free = 20%). The refined atomic model consists of 404 non-hydrogen atoms from the DNA molecule, the atoms from the platinum residue (1 platinum and 2 nitrogen atoms) and 92 water molecules. Special care has been taken about the geometry of the platinum residue. At the end of refinement, the average temperature factor was 18.07 Å² for the nucleic acid atoms. The water molecules have been included in the model with occupancies set to 1.0 and the temperature factors are between 12.5 and 60 Å² (average, 30.25 Å²). The root mean square (r.m.s.) deviation from ideal stereochemistry is 0.007 Å for distances and 0.016 Å for bond angles. The refinement statistics are given in Table 1. Other programs used in this work were from the CCP4 package (40).

RESULTS AND DISCUSSION

We determined the structure of a cisplatin ICL at a high resolution and the quality of the initial map allowed the building of a model which needed little alteration through the refinement of atomic parameters (Fig. 1). This model allows now the description and discussion of the structure at a fine level.

Analysis of the distortions induced by ICL in DNA

The major distortions are located at the level of the adduct and do not extend over the flanking nucleotide residues (Fig. 2). The platinum group which cross-links bases G5* and G15* through their N7 positions, protrudes in the minor groove of the DNA duplex and these N7 atoms located initially in the major groove of B-DNA are now in the minor groove. This striking conformational rearrangement of the molecule results essentially from modifications of the backbone angles P-α-O5', P-ζ-O3' and C5'-γ-C4', belonging to G5* and G15* residues. All the backbone angles are given in Table 2 and the remarkable deviations from standard B-DNA are in bold characters. In the platinated duplex, the heterocyclic O4' moieties of G5* and G15* point towards the 3' direction instead of the 5' direction. Consequently, the O4' of G5* and T7 on one hand and of G15* and G17 on the other hand face each other and their distances are 3.64 and 3.30 Å, respectively. Remarkably, the combination of these three backbone angle values still maintain G5* and G15* phosphate groups oriented so as to keep their O(I) and O(II) atoms pointing towards the solvent (see Fig. 5 for details). The reorganization of the DNA duplex leads also to a permutation of the relative positions of G5* and G15* along the helix axis but, they remain in the *anti* conformation. The nucleotides C6 and C16 (respectively complementary to G15* and G5*, in B-DNA) are extruded from the double helix and exposed to the solvent. The only large alteration of their backbone angles concerns the C3'-ε-O3' angle values (an 80° change which puts the 3' phosphate group in an extrahelical position). All the residues but G5*-C16 and C6-G15* are paired normally. The analysis of the sugar puckers and η inclination angles by *CURVES* (41) indicate a B-like form (Table 2 and Fig. 3a). As a result of the local distortions due to the binding of cisplatin, the helix axis of the cross-linked duplex exhibits a bend of 47° towards the minor groove, the apex of the bending angle being at equal distances from the ends of the molecule. The bend value has been obtained excluding the four central G* and C residues from the computations. This value of 47° is related to the angle defined by the two sets of parallel bases pairs on both sides of the lesion (Fig. 2a, b and c). Furthermore, the value of the

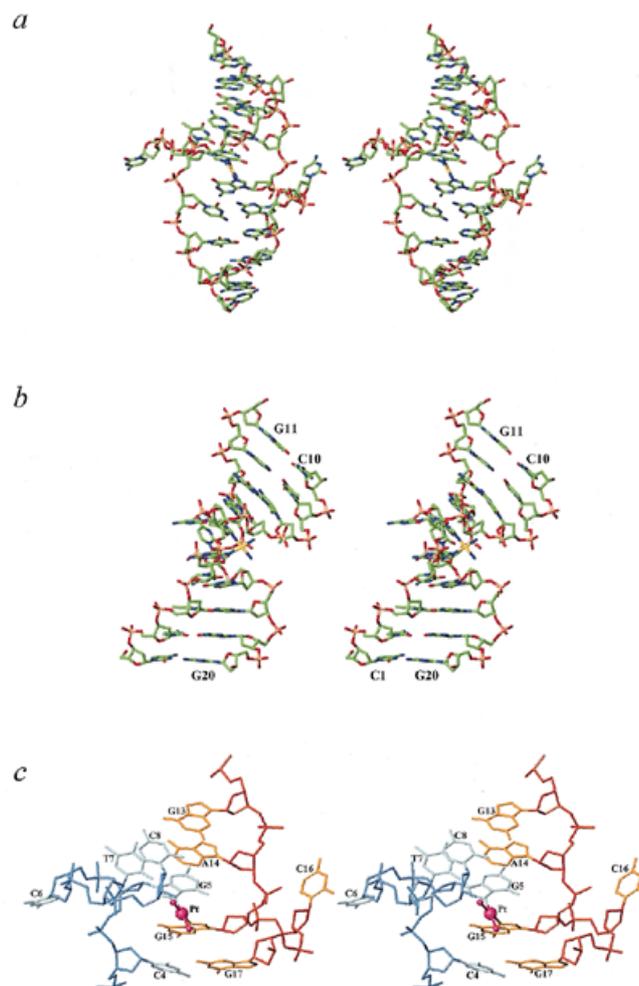


Figure 2. Stereo-views of the atomic model. (a) View showing the minor groove and the bound platinum residue. (b) View after a 90° rotation around a vertical axis. Color code is as Figure 1. (c) Detailed view of the adduct with color code: blue, polypyrimidic strand (pale blue for the bases); red, polypurinic strand backbone (orange for the bases); and purple, platinum residue. Stick representation is from *MOLMOL* (50).

dihedral angle between the mean planes of the two cross-linked guanine rings is 50°, very close to the value of the bending angle of the double helix. This is not a coincidence since G5* and G15* are stacked with A14 and C4, respectively, and are, therefore, parallel to the other Watson–Crick base pairs of the duplex. The minor groove is enlarged to >10 Å (Fig. 3c). In B-DNA, the minor groove width varies between 4 and 6 Å depending on the sequence and assuming a 5.6 Å radius of the phosphate group. The widening of the minor groove is related to a reduction of the inter-base twist angles which results in an helix unwinding (Fig. 3b). Considering the distortions of the double helix, the unwinding measurement between distant base pairs must be precisely defined for comparison between the available models. Indeed, the orientation of both extrahelical cytosine bases is not related to those of the cross-linked guanines and far from being coplanar to them. Hence, the Ω twist angle values given by the program *CURVES* are irrelevant at the level of the lesion

Table 2. Backbone torsion angles α - ζ , glycosyl angles χ , pseudorotation phase angles and amplitudes, and sugar puckers^a

residue	α (deg)	β (deg)	χ (deg)	γ (deg)	δ (deg)	ϵ (deg)	ζ (deg)	phase (deg)	amplitude	pucker
C1	-68	175	-167	46	80	-155	-63	21	37	C3'-endo
C2	-65	-173	-155	56	80	-172	-70	23	35	C3'-endo
T3	-50	166	-139	48	83	-143	-92	22	31	C3'-endo
C4	-69	-149	-98	56	136	-134	-82	169	32	C2'-endo
G5*	85	170	-172	-54	147	-106	89	191	35	C3'-exo
C6	-53	-159	-117	51	135	-82	-55	167	34	C2'-endo
T7	-61	-178	-110	54	138	177	-98	167	33	C2'-endo
C8	-72	166	-114	51	126	-174	-86	146	33	C2'-endo
T9	-60	162	-122	57	88	-175	-88	86	38	O1'-endo
C10	-	-	-134	51	73	-	-	47	42	C4'-exo
G20	-	-	-110	47	139	-	-	162	37	C2'-endo
G19	-74	-164	-179	-175	90	-156	-71	82	34	O1'-endo
A18	-101	73	-85	38	130	-151	-74	143	49	C1'-exo
G17	-77	170	-100	53	148	-163	-81	189	39	C3'-exo
C16	-41	-159	-128	52	145	-86	-63	165	42	C2'-endo
G15*	79	-155	-162	-66	145	-145	103	183	34	C3'-exo
A14	-68	-172	-103	56	138	-125	-78	160	44	C2'-endo
G13	-68	175	-110	42	140	-172	-100	175	34	C2'-endo
A12	-68	162	-95	46	136	-131	-162	157	41	C2'-endo
G11	-67	179	-99	177	150	178	-94	200	37	C3'-exo

^aBackbone torsion angles are O3'--P-(α)-O5'-(β)-C5'-(γ)-C4'-(δ)-C3'-(ϵ)-O3'-(ζ)-P--O5'.

(Fig. 3b). Moreover, the phosphodiester backbones of each strand at the level of the extrahelical cytosines are making right-handed turns, ruling out the use of analytical systems and conventions designed for standard DNA. To measure the relative Ω angle between the base pairs flanking immediately the lesion (C4-G17 and T7-A14), we projected these base pairs on the bisector plane of the dihedral formed by the two pairs, using the *PLOT* option of *TURBO-FRODO*. The Ω angle measured manually on the plot has been compared to the corresponding value in B-DNA. We find an unwinding of 70° for these two base pairs. The bending and unwinding angles of our model are close to those deduced previously from the electrophoretic mobility assays (45 and 80° , respectively; 22,23).

We have also compared our crystallographic model with the two NMR models (Table 3 and Fig. 4). The two DNA sequences used in the NMR studies are different but that of NMR2 is identical to ours. Both NMR models locate the adduct in the minor groove and exhibit a DNA conformation which leads to the extrusion of the two nucleotides C6 and C16 out of the double helix. Application of *CURVES* to the model NMR1 gives a bending angle of 14° and a bend direction close to ours. In contrast, the model NMR2 has a bending angle of 49° but a different bending direction (42° away from ours). The unwinding values for the base pairs adjacent to the adduct, are of 38 and 98° for NMR1 and NMR2, respectively. Considering now the terminal base pairs C1-G20 and C10-G11, the unwinding values are of 70 and 124° for NMR1 and NMR2, respectively (the crystallographic model yielded a value of 110°). At the level of the lesion, the widths of the minor groove given by *CURVES* for the crystallographic model (10.6 \AA) and the two NMR models are about the same.

In the crystallographic structure, the planar character of the platinum coordination is preserved (Fig. 5, Table 4). With the SHEL1 refinement, the four Pt-N bonds are 2.0 \AA long and the bond angles are close to 90° . With SHEL2, the conformation becomes slightly rectangular, with 83.5 and 85.1° for

N7(G5*)-Pt-N7(G15*) and N10-Pt-N20 angles, respectively. N10 and N20 are the two amines of the platinum residue (Table 4). The two other bond angles of the platinum ligand coordination have values of 95.1 and 96.3° (Table 4). These slight differences with respect to a perfect square do not seem distributed at random and may be of chemical significance. As stated in Materials and Methods, the platinum atom as well as the two amines N10 and N20 are resolved, even in the initial map computed at 1.63 \AA resolution. Thus, we consider that the results obtained without restraints for the platinum residue are reliable. Besides, some important distortions occur between the Pt-N7 bonds and the base planes of the cross-linked guanines. The corresponding angles for Pt-N7(G5) and Pt-N7(G15) are 9 and 18° , respectively. Moreover, the platinum atom is displaced by 0.3 and 0.6 \AA from the G5* and G15* base ring planes, respectively, suggesting significant stereochemical constraints. There is only one direct interaction between the platinum residue and a nucleotide not involved in the cross-link. It consists of a hydrogen bond (2.9 \AA) between the ammine group N10 and the O2 of T7 which is not normally involved in the T-A base pairing (Fig. 5). Interestingly, this type of interaction does not exist between the N20 ammine group and the amine N2 of G17. Amine N2 is linked to N20 through water molecules. The other hydrogen bonds involving the platinum residue are with water molecules and will be described below.

Hydration at the platinum site

Water molecules are constituents of nucleic acid structures. They are necessary for stabilization of the double helix and they display a wide variety of patterns which are related to the DNA conformation and to the sequence (42). The structure presented here shows a network of ordered water molecules (Fig. 5). This network forms a cage which surrounds the platinum and the cross-linked guanine residues. Two water molecules (Wa3 and

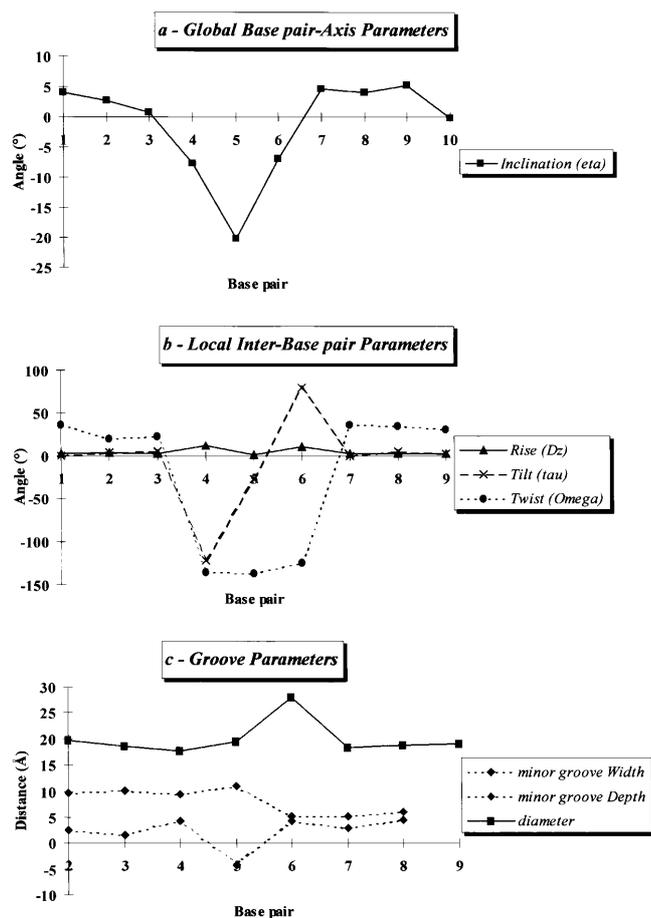


Figure 3. Base pair parameters as a function of the base number of strand 1 as given by the program *CURVES*.

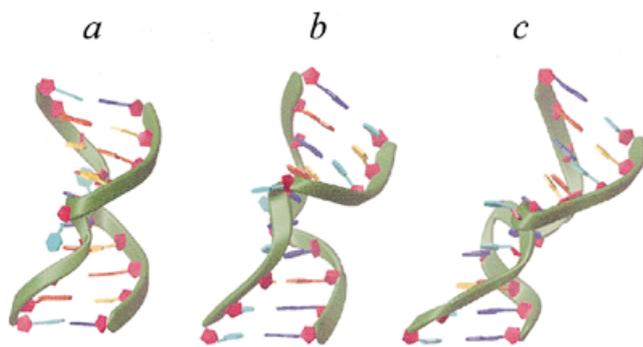


Figure 4. The different available models (*RIBBONS* representation; 51) of the ICL in double-stranded decamers (platinum residue not shown). (a) NMR1 model (23), (b) crystallographic model, (c) NMR2 model (24). The color code for nucleotide bases is: G, purple; C, blue; A, orange; T, yellow, and d-ribose cycles, red.

Wa10) on either side of the square plane of the platinum residue, are located at 3.65 and 3.60 Å from the platinum along its quaternary axis. Each of the oxygen atoms is at hydrogen bond

Table 3. Distortion parameters of cross-linked DNA for the three available models

	NMR1 model	Crystallographic model	NMR2 model
Bend angle (°)	13.7	46.9	49.3
Bend direction* (°)	-53.3	-59.5	-101.1
Global unwinding (°) Base pairs 1-20 and 10-11	70.0	110.0	124.0
Local unwinding (°) Base pairs 4-17 and 7-14	38.0	70.0	98.0
5P to 15P distance (Å)	8.8	15.5	9.2

*Origin of angle is the pseudo-dyad axis *x* of the pair 1-20.

Table 4. Bond distances and angles of the platinum coordination as resulting from refinements SHEL1 and SHEL2

	SHEL1 (with restraints)	SHEL2 (without restraints)
Bond distances (Å) :		
Pt-N7(G5)	2.00	1.98
Pt-N7(G15)	2.01	2.01
Pt-N10	2.00	2.06
Pt-N20	1.99	1.99
Angles (°) :		
N7(G5)-Pt-N7(G15)	86.56	83.50
N7(G5)-Pt-N10	91.78	95.09
N7(G15)-Pt-N20	92.60	96.35
N10-Pt-N20	89.07	85.06

distance (between 2.4 and 3.2 Å) from four or three other hydrogen bond donors or acceptors [Wa3 with Wa9, Wa78, Wa45, O6(G15) on one hand, and Wa10 with Wa67, Wa17 and O6(G5) on the other hand]. N10 is linked to water molecules of the cage Wa17, Wa45, Wa69 and to O2 from T7. N20 is linked to Wa9, Wa69 and Wa12. The cage is made up of nine water molecules immediately surrounding the platinum residue, the two O6 atoms from the cross-linked guanines and the two ammine groups. The cage is linked to G5* and G15* phosphate groups by other water molecules and the distance between the G5* and G15* phosphorus atoms is 15.6 Å. A total of 16 water molecules including the nine of the cage are making contacts with atoms of the adduct. By comparison with both NMR structures, the conformations of the phosphodiester backbones are substantially different at the level of the ICL. Indeed, the water molecules Wa3 and Wa10 are replaced by oxygen atoms from the phosphate groups G5* and G15* with distances between phosphorus G5* and G15* of 8.8 and 9.2 Å for NMR1 and NMR2, respectively (versus 15.6 Å for the crystallographic model). The modeling programs applied to NMR data in the absence of well-ordered water molecules tended to place oxygen atoms of phosphate groups at these precise locations (Fig. 6a and b), probably because of the influence of the electrostatic field around the platinum residue. This shows *a contrario* that water molecules are essential to explain the energy equilibrium of the molecule. As already discussed, the orientation of the phosphate groups G5* and G15* plays an important role in the distortion of the double helix. They are linked by hydrogen bonded water molecules with the cage surrounding the platinum residue and this linkage therefore contributes to maintain their orientation. Furthermore, the cage establishes a link between O6(G5*), O6(G15*) and the platinum residue, contributing to the orientation of these bases, otherwise stacked with the base pairs

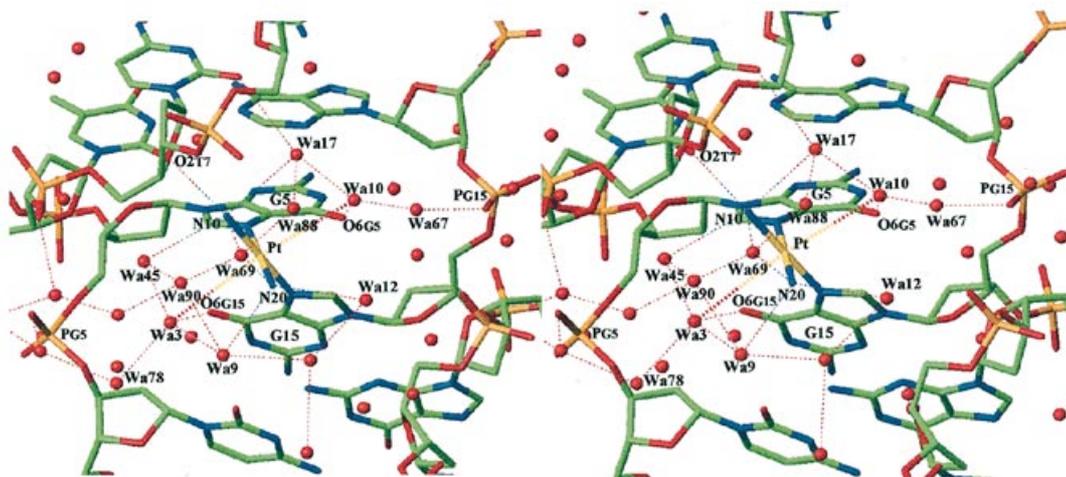


Figure 5. Stereo-view of the cross-linked nucleotides G5* and G15* with the hydrogen bonds between well-ordered water molecules, the cisplatin adduct and between N10 and O2 of T7. Representation program and color code are as in Figure 2a and b.

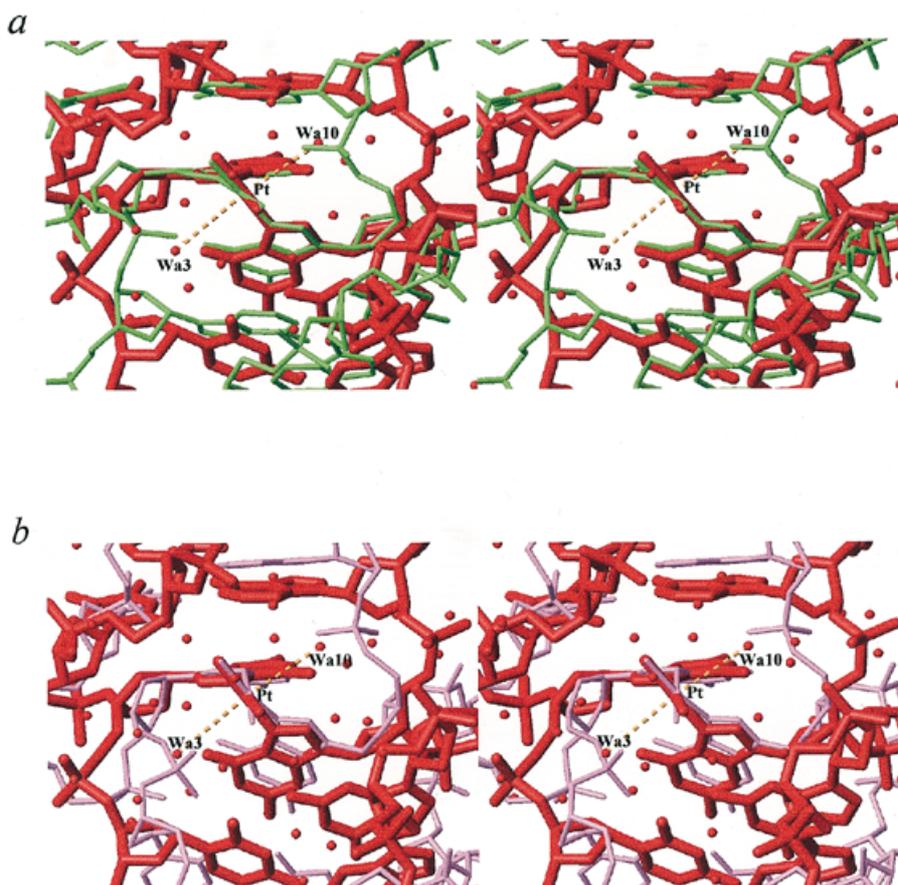


Figure 6. Comparison of the crystallographic model (red) superimposed with (a) NMR1 (purple) and (b) NMR2 (green) models for the same area (as Fig. 5). Stick representation is from *MOLMOL*.

flanking the lesion. Through these interactions with the phosphate groups and the cross-linked guanines, the cage participates to the overall bending of the double helix. The direct or water-mediated interactions between the ammine groups and the bases flanking

the lesion add another contribution to the distortions of the double helix. The two water molecules Wa3 and Wa10 whose oxygen atoms complete an octahedron with the four nitrogen atoms liganded to platinum are favorably located to hydrolyze the Pt-N7 bonds.

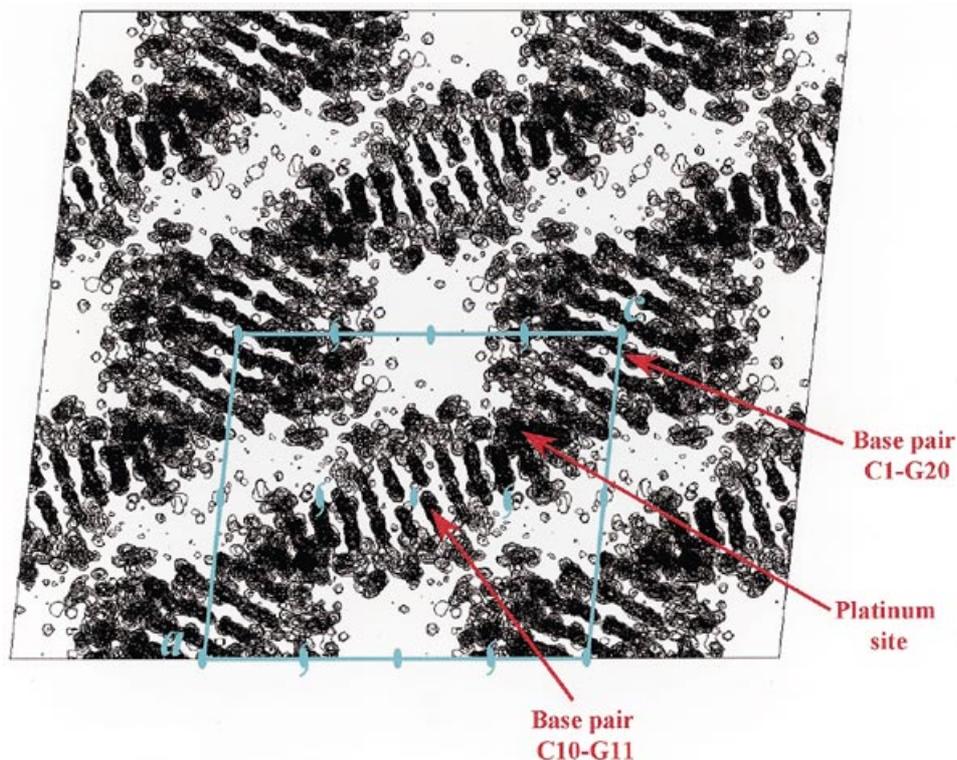


Figure 7. Partial projection ($0 \leq y \leq \frac{1}{2}$) of several crystallographic cells of the initial map as given by *NPO* (40). The unit cell and the symmetry elements are in blue.

Other well-ordered water molecules

The exact number of water molecules which can be located from crystallographic analysis may depend not only on quality and resolution of the data but also on the refinement strategy. This problem is not always addressed in papers devoted to hydration in nucleic acid structures. Recently, the use of MAD crystallographic phases of high quality in structure refinement was able to yield very detailed information about the solvation layer, whatever the chemical nature of the macromolecule (43). Such an approach for the ICL structure is under study with the help of our MAD phases. At the moment, amongst the 92 water molecules so far identified with the classical refinement strategy, 16 are associated to the cage. Most of the other water molecules form chains. One of these chains includes six water molecules and is located in the minor groove along the phosphodiester backbone of the purine residues 17–20. The chain is hydrogen bonded to O6 of guanine residues. In the major and minor grooves, other chains are linked to phosphate groups. One of them which includes seven molecules in the major groove is connected to the N2 atom of the cross-linked guanines. The extrahelical C6 and C16 residues are involved in intermolecular contacts and are surrounded by many well-ordered water molecules. We find an average of four water molecules per residue belonging to the primary layer of hydration, not taking into account the water molecules of the cage.

Crystal packing analysis: evidence of the formation of a triplet between a GC pair and a cytosine extruded from the double helix

The influence of the crystal packing on the conformation of DNA has been already discussed (44). Double-stranded DNA is a

polymorphic molecule and its structure depends on solution conditions. Intermolecular forces in the crystal may shift the equilibrium between several conformations towards one of them. The crystals of the DNA molecule modified by a single cisplatin intrastrand cross-link (14) exhibit two independent molecules in the asymmetric unit and their bending angles have different values. In our case, only one molecule is present in the asymmetric unit. We will discuss the possible influence of the intermolecular interactions due to the packing on the conformation of the molecule.

The projection along the *b* axis ($0 \leq y \leq \frac{1}{2}$) of the map computed with MAD phases and after solvent flattening is illustrated in Figure 7 for several adjacent unit cells. The closest intermolecular contacts are due to end-to-end stacking between bases of contiguous molecules roughly aligned along the crystallographic row $[-1 \ 0 \ 1]$. A single molecule is related to either adjacent molecule along this direction by a 2-fold symmetry axis parallel to *b*. However, the two kinds of contacts at both ends of a single molecule are strikingly different. The terminal pair C10-G11 is stacked with the symmetry-related pair C10'-G11', mimicking the continuity of the double helix. At the other end of the molecule (pair C1-G20) a part of the density belongs to C6 and C16 residues from other molecules located above and underneath along the *b* direction. Indeed, the two cytosine residues extruded from the double helix and belonging to two different molecules obtained by translation (C6'' and C16'') make contacts with the two symmetry related base pairs C1-G20 and C1'-G20'. C6'' is involved in a Hoogsteen base pairing with the pair C1-G20 and C16'' is hydrogen bonded to the phosphate of residue C1 from another molecule. The triplet C1-G20-C6'' is shown in Figure 8a as well as the C16'' residue from another neighboring molecule. On the whole, the contacts at the C1-G20 end involve eight

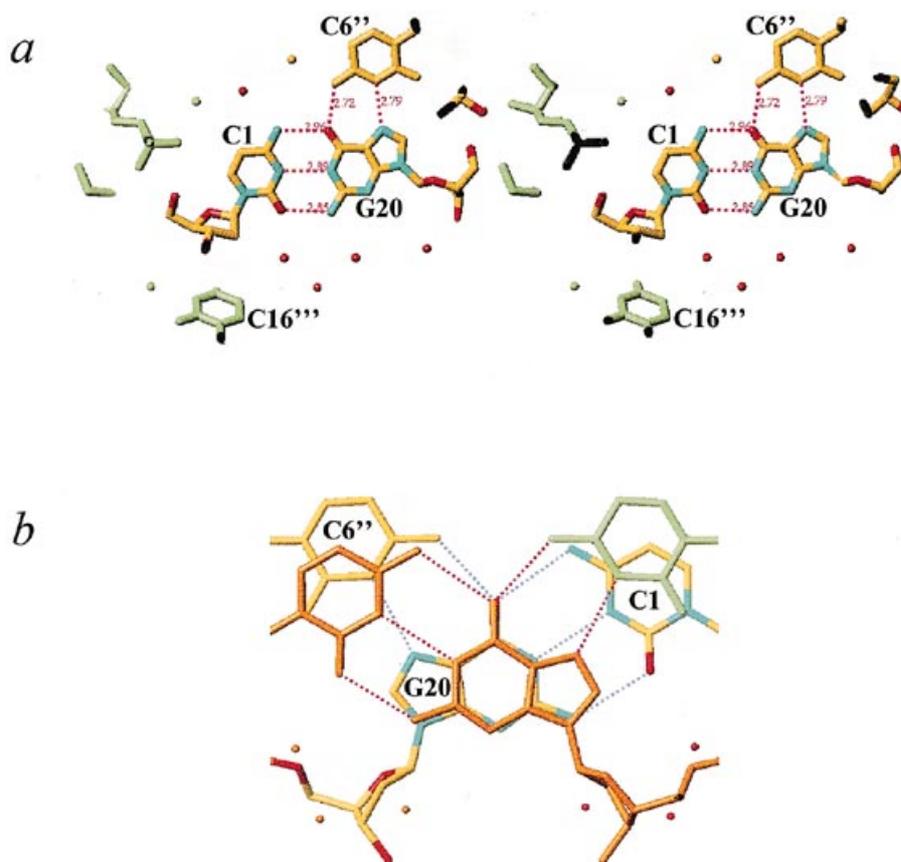


Figure 8. (a) Triplet formed by C1-G20 base pair with C6'' and relative position of the cytosine C16''' which interacts with the phosphate from nucleotide G20. (b) Stacking between the triplet C1-G20-C6'' and the symmetry related one (2-fold axis parallel to the triplets) involving other molecules.

different nucleotides and six different duplexes. The stacking between the two triplets from two symmetrical molecules related by the 2-fold axis is unusual (Fig. 8b). The interaction between two successive triplets should be different in a triple helix where the stacking is constrained by the phosphodiester backbone. The two guanines G20 of the two triplets are partly superimposed whereas the cytosine from one of the pairs C1-G20 is stacked with the extrahelical cytosine C6'' from the symmetrical triplet.

The different types of specific contacts explain the crystal packing in two directions. In one direction, the helix is reconstituted by base pair stacking between successive molecules. Along the *b* axis, the contacts are insured by the two extrahelical cytosines which interact in two different ways. The number of specific intermolecular contacts displayed in this crystal structure explains the high resolution of the diffraction data, the rather low solubility of the molecule in the crystallization conditions and the relatively short crystallization times. The extrahelical nucleotides play a key role in this packing. This result must be compared to previous crystallographic works demonstrating the ability of pending guanines to favor intermolecular triplets and crystallization (45,46). The intermolecular contacts, however, seem of little influence on the conformation of the cross-linked DNA except for the triplet C1-G20-C6''. Stacking between the base pairs from adjacent molecules is unlikely to interfere with the bending, as well as the contacts formed by the extrahelical cytosines which are perpendicular to the bend plane of the ICL. Conversely, the

interactions involving the bases C6'' and C16''' may exert constraints in a direction parallel to the base pair planes and are therefore likely to interfere with the unwinding of the double helix at the C1-G20 base pairs level (110° between terminal base pairs compared to 70° in the vicinity of the lesion).

CONCLUSION

We determined a model of the ICL in double-stranded DNA which only relies on physical information provided by diffraction data and basic stereochemistry. The resolution of the data and the quality of the *ab initio* MAD phases insure that the proposed solution is unique and does not depend on results from other experimental techniques or interaction energy models. A precise description of the distortions induced in DNA by cisplatin ICL has been deduced from the crystallographic data reported here. In particular, it stresses the importance of water molecules in the overall structure of the adduct. The presence of water molecules in the crystallographic structure of *cis*-[Pt(NH₃)₂d{pGpG}] and of a complex between a platinum (IV) derivative with methylcytosinato ligands has already been reported (47,48). However, neither of these structures show the kind of water molecule network observed for the ICL. Our results suggest that some of the chemical and biochemical properties of the ICL are related at least in part, to the hydration of the adduct. Indeed, the presence of the cage of water molecules around the lesion could interfere with the

lability of the cross-link under physiological conditions (49). The bonds between Pt and the N7 atom of G residues within the ICL are spontaneously cleaved, leading to monofunctional adducts which finally rearrange into intrastrand cross-links. In the structure reported here, the two water molecules (Wa3 and Wa10) close to Pt and which complete an octahedron with the two N7G and the two ammines, are in favorable positions for hydrolysis of either N7-Pt bond. This gives a likely explanation to the instability of this lesion and especially to the finding of identical rates of cleavage for the two Pt-N7 bonds. In addition, it is not excluded that stereochemical constraints on the Pt-N7 bonds caused by the displacement of the platinum atom away from the guanine ring planes (0.3 and 0.6 Å for G5* and G15*, respectively) interfere with the ICL lability. The case of the cisplatin 1,2-d(GpG)-intrastrand cross-link which shows instability (12) as well as distortions of the platinum geometry (14) supports this hypothesis.

The particular structure of the ICL raises mechanistic questions about the conformational rearrangement of the DNA molecule. The different steps of the conformational transition from B-DNA to the new structure are unknown. Our results suggest that the water molecules play a role as a driving force in the rearrangement of the platinated duplex. Another more general aspect of the importance of water molecules in DNA is their role in the interaction with proteins. Water molecules are known to be able to mediate the recognition of DNA sequences by proteins (for example in the case of the tryptophan repressor-DNA complex) (29). Repair of cisplatin ICL in the nuclear genome of human cells has been demonstrated (16). The presence of a network of water molecules might be relevant for the mechanism of recognition of the ICL by proteins in addition to the structural parameters of the distorted double helix. Coordinates [SHEL1] and structure factor amplitudes have been deposited in the Nucleic Acid Database, Rutgers University and assigned the NDB accession no. DDJ075.

ACKNOWLEDGEMENTS

We thank Dr Giraud-Panis for helpful comments. This work was supported by grants from l' Association pour la Recherche sur le Cancer, la Ligue Contre le Cancer du Loiret and l'Agence Nationale pour la Recherche sur le SIDA. F.C. is a recipient of a PhD fellowship from the Région Centre.

REFERENCES

- Pratt, W.B., Ruddon, R.W., Ensminger, W.D. and Maybaum, J. (1994) In Pratt, W.B. (ed.), *The Anticancer Drugs*, 2nd Ed. Oxford University Press, NY, pp. 133–139.
- Eastman, A. (1987) *Pharmacol. Ther.*, **34**, 155–166.
- Reedijk, J. (1992) *Inorg. Chim. Acta*, **198**, 873–881.
- Bruhn, S.L., Toney, J.H. and Lippard, S.J. (1990) In Lippard, S.J. (ed.), *Progress in Inorganic Chemistry, Bioorganic Chemistry*. John Wiley & Sons, Inc., New York, pp. 477–516.
- Sip, M. and Leng, M. (1993) In Eckstein, F. and Lilley, D.M.J. (eds), *Nucleic Acids and Molecular Biology*, Vol. 7. Springer Verlag, Berlin and Heidelberg, pp. 1–15.
- Lemaire, M.-A., Schwartz, A., Rahmouni, A.R. and Leng, M. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 1982–1985.
- Hopkins, P.B., Millard, J.T., Woo, J., Weidner, M.F., Kirchner, J.J., Sigmanson, S.T. and Raucher, S. (1991) *Tetrahedron*, **47**, 2475–2489.
- McA' Nulty, M.M. and Lippard, S.J. (1995) In Eckstein, F. and Lilley, D.M.J. (eds), *Nucleic Acids and Molecular Biology*, Vol. 9. Springer Verlag, Berlin and Heidelberg, pp. 264–284.
- Vichi, P., Coin, F., Renaud, J.P., Vermeulen, W., Hoeijmakers, J.H.J., Moras, D. and Egly, J.M. (1997) *EMBO J.*, **16**, 7444–7456.
- Comess, K.M. and Lippard, S.J. (1993) In Neidle, S. and Waring, D. (eds), *Molecular Aspects of Anticancer Drug-DNA Interactions*, Vol. 1. Macmillan Press, London, pp. 134–168.
- Kozelka, J., Fouchet, M.H. and Chottard, J.-C. (1992) *Eur. J. Biochem.*, **205**, 895–906.
- Yang D., van Boom, S.S., Reedijk, J., van Boom, J.H. and Wang, A.H.-J. (1995) *Biochemistry*, **34**, 12912–12920.
- Gelasco, A. and Lippard, S.J. (1998) *Biochemistry*, **37**, 9230–9239.
- Takahara, P.M., Rosenzweig, A.C., Frederick, C.A. and Lippard, S.J. (1995) *Nature*, **377**, 649–652.
- Roberts, J.J. and Friedlos, F. (1987) *Pharmacol. Ther.*, **34**, 15–246.
- Zhen, W., Link, C.J., Jr, O'Connor, P.M., Reed, E., Parker, R., Howell, S.B. and Bohr, V.A. (1992) *Mol. Cell. Biol.*, **12**, 3689–3698.
- Larminat, F., Zhen, W. and Bohr, V.A. (1993) *J. Biol. Chem.*, **268**, 2649–2654.
- Petersen, L.N., Mamenta, E.L., Stevnsner, T., Chaney, S.G. and Bohr, V.A. (1996) *Carcinogenesis*, **17**, 2597–2602.
- Corda, Y., Job, C., Anin, M.F., Leng, M. and Job, C. (1993) *Biochemistry*, **32**, 8582–8588.
- Sip, M., Schwartz, A., Vovelle, F., Ptak, M. and Leng, M. (1992) *Biochemistry*, **31**, 2508–2513.
- Schwartz, A. and Leng, M. (1994) *J. Mol. Biol.*, **236**, 969–974.
- Malinge, J.M., Pérez, C. and Leng, M. (1994) *Nucleic Acids Res.*, **22**, 3834–3839.
- Huang, H., Zhu, L., Reid, B.R., Drobný, G.P. and Hopkins, P.B. (1995) *Science*, **270**, 1842–1845.
- Paquet, F., Pérez, C., Leng, M. (1996) *J. Biomol. Struct. Dyn.*, **14**, 67–77.
- Privé, G.G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M.L. and Dickerson, R.E. (1987) *Science*, **238**, 498–504.
- Schneider, B., Cohen, C. and Berman, H.M. (1992) *Biopolymers*, **32**, 725–750.
- Chalikian, T.V., Plum, G.E., Sarvazyan, A.P. and Breslauer, K.J. (1994) *Biochemistry*, **33**, 8629–8640.
- Hummer, G., Garcia, A.E. and Soumpasis, D.M. (1995) *Biophys. J.*, **68**, 1639–1652.
- Schwabe, J. (1997) *Curr. Opin. Struct. Biol.*, **7**, 126–134.
- Mayer-Jung, C., Moras, D. and Timsit, Y. (1998) *EMBO J.*, **17**, 2709–2718.
- Moy, J.P. (1994) *Nucl. Instr. Meths.*, **A348**, 641–644.
- Evans, P.R. (1993) *Proceedings of CCP4 Study Weekend on Data Collection and Processing*. Daresbury Laboratory, Warrington, UK, pp. 114–122.
- Hendrickson, W.A. (1991) *Science*, **254**, 51–58.
- Leslie, A.G.W. (1990) *Crystallographic Computing*. Oxford University Press, New York, Oxford.
- de la Fortelle, E. and Bricogne, G. (1997) In Sweet, R.M. and Carter, C.W. (eds), *Methods in Enzymology: Macromolecular Crystallography*. Academic Press, NY, pp. 472–494.
- Abrahams, P.J. and Leslie, A.G.W. (1996) *Acta Crystallogr.* **D52**, 34–42.
- Roussel, A. and Cambillau, C. (1991) *Silicon Graphics Geometry Partners Directory*. Silicon Graphics, Mountain View, CA, pp. 77–78.
- Sheldrick, G. SHELX97-Program for Refinement of Crystal Structures. University of Goettingen, Germany.
- Lamzin, V.S. and Wilson, K.S. (1993) *Acta Crystallogr.*, **D49**, 129–147.
- Collaborative Computational Project No. 4 (1994) *Acta Crystallogr.*, **D50**, 760–763.
- Lavery, R. and Sklenar, H. (1988) *J. Biomol. Struct. Dyn.*, **6**, 63–91.
- Berman, H. (1994) *Curr. Opin. Struct. Biol.*, **4**, 345–350.
- Temple Burling, F., Weis, I.W., Flaherty, K.M. and Brunger, A.T. (1996) *Science*, **271**, 72–77.
- Dickerson, R.E., Goodsell, D.S. and Neidle, S. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 3579–3583.
- van Merveelt, L., Vlieghe, D., Dautant, A., Gallois, B., Precigoux, G. and Kennard, O. (1995) *Nature*, **374**, 742–744.
- Vlieghe, D., van Merveelt, L., Dautant, A., Gallois, B., Precigoux, G. and Kennard, O. (1996) *Science*, **273**, 1702–1705.
- Sherman, S.E., Gibson, D., Wang, A.H.-J. and Lippard, S.J. (1985) *Science*, **230**, 412–417.
- Randaccio, L., Zangrando, E., Cesàro, A., Holthenrich, D. and Lippert, B. (1998) *J. Mol. Struct.*, **440**, 221–226.
- Pérez, C., Leng, M. and Malinge, J.M. (1997) *Nucleic Acids Res.*, **25**, 896–903.
- Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.
- Priestle, J.P. (1988) *J. Appl. Crystallogr.*, **21**, 572–576.